

AD-A129 296

OPTIMAL CONTROL OF MARKOV PROCESSES(U) BROWN UNIV
PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMICAL SYSTEMS
W H FLEMING MAR 83 LCDS-83-4 AFOSR-TR-83-0492
AFOSR-81-0116

1/1

UNCLASSIFIED

F/G 12/1

NL

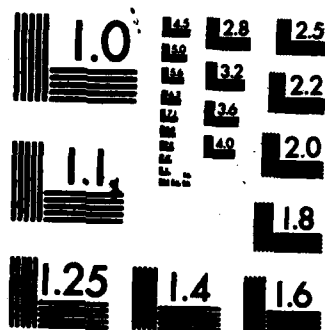


END

DATE

FILED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A129

DTIC FILE COPY

Lefschetz Center for Dynamical Systems

Approved for public
distribution under

OPTIMAL CONTROL OF MARKOV PROCESSES

by

Wendell H. Fleming

March 1983

LCDS #83-4

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)

NOTICE OF TRANSMITTAL TO DTIC

This technical report has been reviewed and is
approved for public release IAW AFR 190-12.

Distribution is unlimited.

MATTHEW J. KERPER

Chief, Technical Information Division

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83 - 0492	2. GOVT ACCESSION NO. A129296	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) OPTIMAL CONTROL OF MARKOV PROCESSES		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
7. AUTHOR(s) Wendell H. Fleming		6. PERFORMING ORG. REPORT NUMBER LCDS #83-4
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lefschetz Center for Dynamical Systems Division of Applied Mathematics, Brown University, Providence RI 02912		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0116
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical & Information Sciences Directorate Air Force Office of Scientific Research Bolling AFB DC 20332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A4
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE MAR 83
		13. NUMBER OF PAGES 22
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Invited Plenary Paper, International Congress of Mathematicians, Warsaw, August 1983.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The purpose of this article is to give an overview of some recent developments in optimal stochastic control theory. The field has expanded a great deal during the last 20 years. It is not possible in this overview to go deeply into any topic, and a number of interesting topics have been omitted entirely. The development of stochastic control theory has depended on parallel advances in the theory of stochastic processes and on certain topics in partial differential equations. On the probabilistic side one can mention (CONTINUED)		

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

83 06 10 028

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

ITEM #20, CONTINUED: decomposition and representation theorems for semi-martingales, formulas for absolutely continuous change of probability measure (e.g., the Girsanov formula), and the study of Ito-sense stochastic differential equations with discontinuous coefficients. It seems fair to say that these developments in stochastic processes were in turn to an extent influenced by their applications in stochastic control. For controlled Markov diffusion processes, there is a direct connection with certain nonlinear partial differential equations via the dynamic programming equation. These equations are of second order, elliptic or parabolic, and possibly degenerate. Stochastic control gives a way to represent their solutions probabilistically. There is an unforeseen connection with differential geometry via the Monge-Ampere equation.

x Sub t

Broadly speaking, stochastic control theory deals with models of systems whose evolution is affected both by certain random influences and also by certain inputs chosen by a "controller". The authors are concerned here only with state-space formulations of control problems in continuous time. Moreover, the authors consider only markovian control problems in which the state x_t of the process being controlled is Markov provided the controller follows a Markov control policy. The authors shall not discuss at all the extensive engineering literature on input-output formulations particularly for linear system models, see Astrom [1].

They

u Sub t

The authors shall mainly discuss the case of continuously acting control, in which at each time t a control u_t is applied to the system. However, in #8 the authors briefly mention impulsive control problems, in which control is applied only at discrete time instants. In optimal stochastic control theory the goal is to minimize (or maximize) some criterion depending on the states x_t and controls u_t during some finite or infinite time interval. In #2 the authors formulate a class of optimal control problems for Markov processes, with criterion (2.2) to be minimized. The distinction between problems in which x_t is known to the controller, and problems with partial observations is made there. When x_t is known, the dynamic programming method can be used. In principle, this method leads directly to an optimal Markov control policy, although it rarely gives the optimal policy explicitly. In #3, both analytical and probabilistic approaches are indicated. Associated with dynamic programming is the Nisio nonlinear semigroup (#4). In #5 a logarithmic transformation is applied to positive solutions of the backward equation of a Markov process. There results a controlled Markov process, leading to connections between stochastic control and such topics as stochastic mechanics, large deviations and nonlinear filtering. The case of controlled, partially observed processes is mentioned in #7, along with adaptive control of Markov processes. Finally, in #9, the authors indicate a few of the various difficulties encountered in seeking to implement in engineering applications the mathematically sophisticated results of the theory, and mention some newer areas of application.

UNCLASSIFIED

OPTIMAL CONTROL OF MARKOV PROCESSES*

Wendell H. Fleming

Invited Plenary Paper
International Congress of Mathematicians
Warsaw
August 1983

* This research has been supported in part by the Air Force Office of Scientific Research under contract #AFOSR-81-0116 and in part by the National Science Foundation under contract #MCS-8121940.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



OPTIMAL CONTROL OF MARKOV PROCESSES

Wendell H. Fleming

1. Introduction. The purpose of this article is to give an overview of some recent developments in optimal stochastic control theory. The field has expanded a great deal during the last 20 years. It is not possible in this overview to go deeply into any topic, and a number of interesting topics have been omitted entirely. The list of References includes several books, conference proceedings and survey articles.

The development of stochastic control theory has depended on parallel advances in the theory of stochastic processes and on certain topics in partial differential equations. On the probabilistic side one can mention decomposition and representation theorems for semimartingales, formulas for absolutely continuous change of probability measure (e.g. the Girsanov formula), and the study of Ito-sense stochastic differential equations with discontinuous coefficients. It seems fair to say that these developments in stochastic processes were in turn to an extent influenced by their applications in stochastic control. For controlled Markov diffusion processes, there is a direct connection with certain nonlinear partial differential equations via the dynamic programming equation. These equations are of second order, elliptic or parabolic, and possibly degenerate.

Stochastic control gives a way to represent their solutions probabilistically. There is an unforeseen connection with differential geometry via the Monge-Ampere equation.

Broadly speaking, stochastic control theory deals with models of systems whose evolution is affected both by certain random influences and also by certain inputs chosen by a "controller". We are concerned here only with state-space formulations of control problems in continuous time. Moreover, we consider only markovian control problems in which the state x_t of the process being controlled is Markov provided the controller follows a Markov control policy. We shall not discuss at all the extensive engineering literature on input-output formulations particularly for linear system models, see Åström [1].

We shall mainly discuss the case of continuously acting control, in which at each time t a control u_t is applied to the system. However, in §8 we briefly mention impulsive control problems, in which control is applied only at discrete time instants. In optimal stochastic control theory the goal is to minimize (or maximize) some criterion depending on the states x_t and controls u_t during some finite or infinite time interval. In §2 we formulate a class of optimal control problems for Markov processes, with criterion (2.2) to be minimized. The distinction between problems in which x_t is known to the controller, and problems with partial observations is made there. When x_t is known, the dynamic programming method can be used. In principle, this method leads directly to an optimal Markov control policy, although it rarely gives the optimal policy explicitly. In §3, both analytical

and probabilistic approaches are indicated. Associated with dynamic programming is the Nisio nonlinear semigroup (§4). In §5 we discuss methods of approximate solution and special problems. In §6 a logarithmic transformation is applied to positive solutions of the backward equation of a Markov process. There results a controlled Markov process, leading to connections between stochastic control and such topics as stochastic mechanics, large deviations and nonlinear filtering. The case of controlled, partially observed processes is mentioned in §7, along with adaptive control of Markov processes. Finally in §9 we indicate a few of the various difficulties encountered in seeking to implement in engineering applications the mathematically sophisticated results of the theory, and mention some newer areas of application.

2. Controlled Markov processes. We consider optimal stochastic control problems of the following kind. We are given metric spaces Σ , U called the state space and control space, respectively. For each fixed $u \in U$ there is a linear operator L^u which generates a Markov, Feller process with state space Σ . The domain of L^u contains, for each $u \in U$, a set D dense in the space $C(\Sigma)$ of bounded uniformly continuous functions on Σ . The state and control processes x_t, u_t are defined on some probability space (Ω, \mathcal{F}, P) . The Σ -valued process x_t is adapted to some increasing family of σ -algebras $\mathcal{F}_t \subset \mathcal{F}$, and the trajectories x are right continuous. The U -valued process u_t is predictable with respect to an increasing

family of σ -algebras $\mathcal{G}_t \subset \mathcal{F}_t$. The σ -algebra \mathcal{G}_t describes in a measure theoretic way the information available to the controller at time t .

The processes (x_t, u_t) are related by the requirement that

$$(2.1) \quad Mg(t) = g(x_t) - g(x_0) - \int_0^t L^u_s g(x_s) ds$$

is a (\mathcal{F}_t, P) martingale for every $g \in D$. We consider a fixed, finite time interval $0 \leq t \leq T$, and the objective to minimize a criterion of the form of an expectation

$$(2.2) \quad J = E\left\{\int_0^T k(x_t, u_t) dt + G(x_T)\right\}.$$

Example 1. Controlled finite-state Markov chain, with $\Sigma = \{1, 2, \dots, N\}$. In this case L^u is identified with the infinitesimal matrix (q_{ij}^u) of the chain. When the control u_t is applied, the jumping rate of x_t from state i to j is $q_{ij}^{u_t}$.

Example 2. Controlled diffusion process with $\Sigma = \mathbb{R}^n$,

$$(2.3) \quad x_t = x_0 + \int_0^t f(x_s, u_s) ds + \int_0^t \sigma(x_s, u_s) dw_s,$$

with w_t a brownian motion (of some dimension d) independent of the initial state x_0 . In this case

$$(2.4) \quad L^u = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x, u) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n f_i(x, u) \frac{\partial}{\partial x_i}$$

with $a = \sigma \sigma'$ and $D = \{g : g, g_{x_i}, g_{x_i x_j} \in C(\mathbb{R}^n), i, j = 1, \dots, n\}$.

The diffusion is called nondegenerate if the eigenvalues of $a(x, u)$

are bounded below by $c > 0$.

Further assumptions, which vary from author to author in the literature, need to be made. To avoid undue complication, in the discussion to follow we take a compact control space U , and $k(x,u)$, $G(x)$ bounded, uniformly continuous. In (2.3), $f(x,u)$, $\sigma(x,u)$ are bounded and as smooth as necessary. The σ -algebras $\mathcal{F}_t, \mathcal{G}_t$ are right continuous and completed.

If x_t is \mathcal{G}_t -measurable, then the controller can observe the state x_t . In this case, one may as well take $\mathcal{G}_t = \mathcal{F}_t$ and known initial state x_0 . This is the situation in §'s 3-6 to follow. If (2.1) holds, we call

$$\sigma = (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}, x, u.)$$

an admissible system for the control problem with completely observed states.

A Markov control policy is a Borel measurable function from $[0, T] \times \Sigma$ into U . An admissible system σ is obtained via a Markov control policy \underline{u} if

$$(2.5) \quad u_t = \underline{u}(t, x_t^-).$$

Given \underline{u} and $x_0 \in \Sigma$, one would like to know whether a corresponding admissible system exists, with x_t a Markov process. Under sufficiently strong restrictions this is well known. For instance, in case of controlled diffusions a Lipschitz condition on $\underline{u}(t, x)$ would

imply the classical Ito conditions. For nondegenerate controlled diffusions, existence follows from Krylov [8, p.87] for any bounded \underline{u} . The Markov property of x_t can be obtained under stronger hypotheses. For instance, for nondegenerate diffusions it holds if in (2.3) $\sigma = \sigma(x)$. A martingale method for obtaining the Markov property is to show that the probability distribution $P_{x_0}^u$ of the state trajectory x_t is unique and depends continuously on the initial state x_0 [59]. In general x_t is only a weak-sense solution to (2.3), since neither the probability space nor the brownian motion w_t are given in advance. However, in the nondegenerate case with $\sigma = \sigma(x)$ a result of Veretennikov [62] gives a strong solution.

3. Dynamic programming. The dynamic programming approach to the control problem with completely observed states x_t can be described in a purely formal way, as follows. For initial state $x_0 \in \Sigma$ and admissible system σ , write $J = J(T, x_0, \sigma)$ in (2.2). Let

$$(3.1) \quad W(T, x_0) = \inf_{\sigma} J(T, x_0, \sigma).$$

Formal reasoning indicates that $W(T, x)$ should satisfy the dynamic programming equation

$$(3.2) \quad \frac{\partial W}{\partial T} = AW, \quad T > 0,$$

with initial data $W(0, x) = G(x)$, where

$$(3.3) \quad Ag(x) = \min_{u \in U} [L^u g(x) + k(x, u)].$$

Formally, an optimal Markov policy \underline{u}^* is found by requiring $\underline{u}^*(t, x)$ to minimize $L^u W(T - t, x) + k(x, u)$ among all $u \in U$. Instead of the finite time control problem, control until x_t exits a given open set $O \subset \Sigma$ can be considered. In that case the dynamic programming equation becomes the autonomous form of (3.2) in O , with $W(x) = G(x)$ for $x \in \partial O$. There are also autonomous dynamic programming equations associated with the infinite time control problem, with discounted cost or average cost per unit time criteria to be minimized.

In the rigorous mathematical treatment of dynamic programming there is one easy result, the so-called Verification Theorem [7, p. 159]. Roughly speaking, it states that if $W(T, x)$ satisfying (3.2) with the initial data and the associated Markov policy \underline{u}^* are both "sufficiently regular", then \underline{u}^* is indeed optimal and $W(T, x)$ is the minimum performance in (3.1). The Verification Theorem is used to obtain explicit solutions, in those cases where such a solution is known. Much more difficult are the questions of existence of sufficiently regular W and \underline{u}^* , and there is a large literature dealing with various aspects of them. One approach is analytical with the stochastic interpretation made afterward. In this approach, existence of solutions to the dynamic programming equation and their regularity properties are studied, using non-probabilistic methods. It is then proved that optimal (or at least ϵ -optimal) Markov control policies exist. A second approach is probabilistic. In this approach, one starts with the minimum cost function W in (3.1)

and develops stochastic counterparts to the dynamic programming conditions for a minimum. A third approach is to consider an associated nonlinear semigroup (§4). While this approach leads to fewer technical difficulties than either of the other two, it also leads to weaker results.

For controlled diffusions the analytical approach is remarkably well developed. See Krylov [8], Lions [45]. In the nondegenerate case the dynamic programming equation is a second order nonlinear partial differential equation of parabolic type also called a Hamilton-Jacobi-Bellman equation. In various other formulations, with x_t controlled for all time $t \geq 0$ or until exit from an open set O , the Hamilton-Jacobi-Bellman equation is elliptic rather than parabolic. Under reasonable assumptions the problem, the solution S has generalized second derivatives which are locally bounded. In the elliptic case a deeper regularity result of Evans [26][60] gives a classical solution. In the degenerate case W is less regular with locally bounded generalized first derivatives W_{x_i} . The dynamic programming equation (3.2), suitably interpreted in terms of Schwartz distributions, still holds [8][45]. For the case of controlled jump Markov processes, results on existence, uniqueness and regularity of solutions to (3.2) were obtained by Pragarauskas [52].

A large class of nonlinear elliptic or parabolic equations, satisfying appropriate convexity conditions, can be represented as Hamilton-Jacobi-Bellman equations. As Gaveau [35] pointed out, the Monge-Ampere equation has such a representation.

In the probabilistic approach, the starting point is to rewrite the dynamic programming principle in the following martingale form. Given an admissible system σ let

$$m_t = \int_0^t k(x_s, u_s) ds + W(T - t, x_t).$$

Then m_t is a (\mathcal{F}_t, P) submartingale, and σ is optimal if and only if m_t is a (\mathcal{F}_t, P) martingale. With the aid of the Doob-Meyer decomposition for submartingales and some martingale representation theorems, conditions for optimality are obtained. See Bismut [21], Davis [16], Elliott [25], El Karoui [5]. These conditions are probabilistic counterparts of those expressed analytically by the dynamic programming equation (3.2). With the probabilistic approach difficult questions of regularity of solutions to (3.2) are avoided. The probabilistic techniques give results about existence of optimal Markov policies [21], [5, p. 218]. These methods also give conditions for a minimum for optimal control under partial observations.

A different kind of Markovian control problem for diffusions, in which the control acts only on the boundary of a region $\mathcal{O} \subset \mathbb{R}^n$ was considered by Vermes [61].

4. The Nisio nonlinear semigroup. The dynamic programming principle can be restated in another form, in terms of a semigroup of nonlinear operators. In purely formal way, this is done as follows. In (2.2) we fix k but consider various G . We rewrite the infimum in (3.1) as $W(T, x) = S_T G(x)$. The dynamic programming principle is

formally equivalent to the semigroup property

$$(4.1) \quad S_{T_1+T_2} = S_{T_1} \circ S_{T_2}$$

of the family $\{S_T\}$ of nonlinear operators. In addition, for "sufficiently regular" G , one should have

$$(4.2) \quad \frac{d}{dT} S_T G|_{T=0} = AG.$$

This formal procedure was put on a rigorous basis by Nisio [10], who obtained $\{S_T\}$ as a semigroup on the space $C(\Sigma)$ and showed under some mild additional conditions that (4.2) holds for $G \in D$ (notation of §2). Equations (4.1), (4.2) would imply the dynamic programming equation (3.2) if we knew that $W(T, \cdot) = S_T G$ is sufficiently regular (in particular, if S_T maps D into D .) However, W does not generally have the desired regularity. In such instances (4.2) is a kind of weaker substitute for (3.2).

Nisio's treatment is analytical. She obtains S_T as the lower envelope of the family of linear semigroups S_T^u , where for constant control $u \in U$ the generator of S_T^u coincides on D with the operator $L^u + k(\cdot, u)$. A stochastic treatment of the Nisio semigroup is given in Bensoussan-Lions [2], and a uniqueness result in case of nondegenerate diffusions in Nisio [51]. El Karoui, Lepeltier, and Marchal [24] used another procedure, and obtained a nonlinear semigroup on a larger space of bounded functions G which are measurable in a suitable sense.

5. Explicit and approximate solutions. In a few instances the dynamic programming equation (3.2) can be solved explicitly. Examples are the well known stochastic linear regulator and Merton's optimal portfolio selection problem [7, pp. 160,165]. For other special problems the solution can be reduced to a free boundary problem. The boundaries to be determined separate regions where some control constraint holds or not. See for example Karatzas-Benes [40].

When a solution cannot be found by special methods, one can seek an approximate solution to (3.2). One class of approximate methods involve discretizations of (3.2). Among such methods the algorithm of Kushner [9] has a natural stochastic control interpretation. The difference equations associated with the algorithm correspond to the dynamic programming equation for an approximating controlled Markov chain. For the special case of controlled one-dimensional diffusions, Borkar and Varaiya [22] used a procedure in which piece-wise-constant approximating Markov control policies are allowed.

Other results give approximate solutions to (3.2) when the state process x_t is a nearly-deterministic controlled diffusion. In (2.3) let $\sigma = \epsilon^{1/2} \sigma_0$. The solution is sought in the form of an asymptotic series in ϵ . In [27] this is done by expanding the solution $W^\epsilon(T,x)$ in an asymptotic series. The expansion is valid in regions where the solution $W^0(T,x)$ of the corresponding Hamilton-Jacobi equation is smooth. In [20] Bensoussan obtains an asymptotic expansion, using a stochastic maximum principle instead of (3.2).

6. A logarithmic transformation. Consider a linear operator of the form $L + V(x)$, where L is the generator of a Markov process ξ_t with state space Σ . The initial value problem

$$(6.1) \quad \frac{d\phi}{dt} = L\phi + V(x)\phi$$

with data $\phi(0,x) = \phi(x)$ has a probabilistic solution by a well known formula of Feynman-Kac type. For positive solutions of (6.1) another probabilistic representation for $\phi(T,x)$ can often be found in the following way. The logarithmic transformation $I = -\log \phi$ changes (1.1) into the nonlinear equation

$$(6.2) \quad \frac{dI}{dt} = H(I) - V(x),$$

$$(6.3) \quad H(I) = -e^I L(e^{-I}).$$

If one can find a control problem of the kind in §2 such that

$$(6.4) \quad H(I) = \min_{u \in U} [L^u I + k(x,u)],$$

then (6.2) is the dynamic programming equation (3.2). The stochastic control interpretation of $I(T,x)$ is as the minimum of the criterion J in (2.2). Thus, in (3.1) we have $W = I$. For a nondegenerate diffusion obeying the stochastic differential equation

$$(6.5) \quad d\xi_t = b(\xi_t)dt + \sigma(\xi_t)d\omega_t,$$

a Markov control policy $\underline{u}(t,x)$ changes the generator L to $L^{\underline{u}}$, corresponding to change of drift from $b(x)$ to $\underline{u}(t,x)$ in (6.5).

In (2.2) one takes

$$k(x,u) = \frac{1}{2}(b(x) - u)' a^{-1}(x)(b(x) - u),$$

with $a = \sigma\sigma'$. An appropriate control problem for the case of ε_t a jump Markov process is described in [31], and for a general class of Markov ε_t in Sheu's thesis [58]. The change of generator from L to L^u corresponds to a change of probability measure. It was pointed out by M. Day that this change of measure results by conditioning with respect to $\phi(x_T)$. See [31,(4.5)].

In case $L = \frac{1}{2}\Delta$, corresponding to ε_t a brownian motion (6.1) is the heat equation with a potential term. The stochastic control interpretation of $S = -\log \phi$ is as least average action. Upon rescaling, taking $L = \frac{\lambda}{2}\Delta$ and replacing V by $\lambda^{-1}V$, the usual least action is obtained as a "classical mechanical limit" as $\lambda \rightarrow 0$ [29]. The heat equation with potential is the "imaginary time" analogue of the Schrödinger equation of quantum mechanics. There is an intriguing connection between stochastic control and the Schrödinger equation, whose implications are not as yet well understood [36]. This work is in the framework of Nelson's stochastic mechanics. An apparently different theory of "stochastic mechanics" was developed by Bismut [4].

Holland [39] gave a stochastic control interpretation of the dominant eigenvalue of the Schrödinger equation as minimum mean total energy of a particle in equilibrium. The approach was again based on a logarithmic transformation and subsequently led to Sheu's treatment [58]

of the Donsker-Varadhan formula for the dominant eigenvalue of the operator $L + V$ appearing in (6.1).

The Venttsel-Freidlin theory of large deviations deals with asymptotic probabilities of rare events associated with nearly deterministic Markov processes. The logarithmic transform gives another approach to results of this kind. As an illustration we consider the problem of exit from an open set $D \subset E$ during the time interval $0 \leq t \leq T$. Let x_t^ϵ be a Markov process tending to a deterministic limit x_t^0 as $\epsilon \rightarrow 0$. Let $I^\epsilon = -\epsilon \log P_x(\tau^\epsilon \leq T)$, where τ^ϵ is the exit time of x_t^ϵ from D . Under various assumptions (including a suitable scaling of ϵ), I^ϵ tends to a limit I^0 , where $I^0(T, x)$ is the minimum of a certain "action functional" among curves starting at $x \in D$ and leaving D by time T . In the stochastic control approach $I^\epsilon(T, x)$ is the minimum performance in a corresponding stochastic control problem [28] [31] [58]. In this approach a minimum principle is associated with the large deviation problem for $\epsilon > 0$, not just in the limit as $\epsilon \rightarrow 0$.

In [32], the logarithmic transformation was applied to solutions to the pathwise equation of nonlinear filtering, making a connection between filtering and stochastic control.

7: Partial observations; adaptive control. The states x_t of a stochastic system often cannot in practice be measured directly, or perhaps can only be measured with random errors. This has led to an extensive literature on nonlinear filtering and on optimal control under

partial observations. For controlled diffusions, a standard model is to take state dynamics (2.3) and an observation process y_t governed by

$$(7.1) \quad y_t = \int_0^t h(x_s) ds + W_t,$$

with W a brownian motion independent of w . The information available to the controller at time t is usually assumed to be described by the σ -algebra \mathcal{G}_t generated by observations y_s for $s \leq t$. However, existence of optimal controls has been proved only with a somewhat wider class of admissible controls than those adapted to this family $\{\mathcal{G}_t\}$.

Several good survey articles on controlled partially observed diffusions have recently appeared [15][16][17]. Hence, we shall not try to summarize the various results here. In studying partially observed control problems it is useful to introduce an auxiliary "separated" control problem. In the separated problem the role of "state" process is taken by a measure-valued stochastic process σ_t [34]. The measure σ_t represents an unnormalized conditional distribution of x_t given observations and controls $y_s, u_s, 0 \leq s \leq t$. A nonlinear semigroup for the controlled, measure-valued process σ_t has been constructed [19][30][33]. Among other recent work, we mention that of Rishel [53] on partially observed jump processes, and of Mazzitello-Spirglas [49] on impulsive control under partial information.

Adaptive control. In adaptive control the objective is the simultaneous control and identification of unknown system parameters. Common techniques in discrete-time adaptive control involve sequential techniques, based on maximum likelihood or least squares, for updating estimates of unknown parameters. In the context of adaptive control of Markov chains see the pioneering work of Mandl [48], also Borkar-Varaiya [23], Kumar-Lin [41]. Another (Bayesian) viewpoint is to treat adaptive control of Markov processes as a special case of stochastic control under partial observations. This is done by simply regarding the unknown parameters as additional (nontime-varying) components of the system state. From a practical standpoint this approach encounters well known difficulties, in that effective solutions to partially observed stochastic problems are difficult to obtain. Nevertheless, special cases in which the problem becomes finite dimensional have been treated by Hijab [38] and Rishel [54].

8. Impulse control; problems with switching costs. In impulse control problems the control actions are taken at discrete (random) time instance, and each control action leads to an instantaneous change in the state x_t . Typical impulse control problems are those of stock inventory management, in which a control action is to reorder with immediate delivery of the order.

The analytic treatment of impulse control was initiated and developed systematically by Bensoussan and Lions [3], with emphasis on the control

of nondegenerate diffusions. The dynamic programming equation is replaced by a set of inequalities which take the form of a quasivariational inequality. For the case of degenerate diffusions see Menaldi [50], and for impulsive control for Markov Feller process see Robin [55][56]. Lepeltier-Marchal [43] gave a probabilistic treatment.

Another class of stochastic control problems of recent interest are those in which control actions are taken at discrete time instants, with no instantaneous change in x_t but with a cost of switching control actions. Such problems arise in the theory of controlled queues (see Sheng [57]) and in control of energy generating systems under uncertain demand. The analytical treatment again is to reduce the problem to a quasivariational inequality. See Lenhart-Belbas [42], Liao [44].

9. Applications. Optimal stochastic control theory was initially motivated by problems of control of physical devices. More recent influences have come from management science, economics, and information systems. Until now, the impact on engineering practice of much of the sophisticated mathematical theory has been small. The stochastic linear regulator is a standard tool, because the optimal Markov control policies turn out to be linear in the state x . If the Markov policy is nonlinear, it is difficult to implement. Moreover, other issues may be considered in practice more important than optimality of system performance as predicted by the stochastic control model. The model is generally a simplification of nature, through linearizations, reductions of dimensionality, assumptions that noises are white, etc. A control

which performs well (even optimally) according to the model may behave poorly in a real control system. The question of robustness of controls with respect to unmodelled system dynamics is of current interest in the engineering control literature. See for example [63]. A different sort of question is that of stochastic controllability [64].

We conclude by mentioning two novel applications of stochastic control. One is Arrow's model of exploration consumption, and pricing of a randomly distributed natural resource. This model was analyzed in detail by Hagan-Caflisch-Keller [37]. They determined approximately the free boundary between portions of the state space where new exploration should or should not be undertaken.

Ludwig and associates have applied a stochastic control method to fishery management problems [47]. The fishery resource is controlled through the rate at which fish are harvested. This work has an important statistical aspect as well as the control aspect, since errors in measuring unknown parameters in the fishery model can be important.

References

Books and Monographs

1. K. Åström, Introduction to Stochastic Control Theory, Academic Press, 1970.
2. A. Bensoussan and J.L. Lions, Applications des Inéquations Variationnelles en Contrôle Stochastique, Dunod, Paris, 1978.
3. A. Bensoussan and J.L. Lions, Contrôle Impulsionnel et Inéquations Quasi-Variationnelles, Dunod, Paris, 1981.
4. J.-M. Bismut, Mécanique Aléatoire, Springer Lecture Notes in Math. No. 866, 1981.
5. N. El Karoui, Les Aspects Probabilistes du Contrôle Stochastique, pp. 73-238 of Springer Lecture Notes in Math. No. 876, 1981. École d'Été de Probabilités de Saint-Fleur IX-1979.
6. R.J. Elliott, Stochastic Calculus and Applications, Springer Verlag 1982.
7. W.H. Fleming and R.W. Rishel, Deterministic and Stochastic Optimal Control, Springer-Verlag, 1975.
8. N.V. Krylov, Controlled Diffusion Processes, Springer-Verlag, 1980.
9. H.J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, 1977.
10. M. Nisio, Lectures on Stochastic Control Theory, Indian Statistical Institute Lecture Notes Series No. 9, 1981.

Conference Proceedings

11. M. Kohlmann and W. Vogel (eds.) Stochastic Control Theory and Stochastic Differential Systems, Springer Lecture Notes in Control and Info. Sci. No. 16, 1979.
12. W.H. Fleming and L.E. Gorostiza (eds.) Advances in Filtering and Optimal Stochastic Control, Springer Lecture Notes in Control and Info. Sci. No. 42, 1982.
13. M. Kohlmann and R. Christensen (eds.), Stochastic Differential Systems, Springer Lecture Notes in Control and Info. Sci. No. 43, 1982.

Survey Articles

14. J.M. Bismut, An introductory approach to duality in optimal stochastic control, SIAM Review 20 (1978).
15. N. Christopeit and M. Kohlmann, Some recent results on the control of partially observable stochastic systems, in ref. [13] pp. 251-275.
16. M.H.A. Davis, Martingale methods in stochastic control, in ref. [11] pp. 85-117.
17. M.H.A. Davis, Some current issues in stochastic control theory, Stochastics, to appear.
18. P.-L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations, Proc. 1982 Internat. Congress of Math.

Other Articles

19. A. Bensoussan, in Ref. [12] pp. 38-53.
20. A. Bensoussan, in Ref. [13] pp. 216-228.
21. J.M. Bismut, Memoirs Amer. Math. Soc. No. 167, 1976.
22. V. Borkar and P. Varaiya, IEEE Trans. Auto. Control AC-26 (1981), pp. 466-470.
23. V. Borkar and P. Varaiya, SIAM J. Control Optimiz. 20 (1982), pp. 470-489.
24. N. El Karoui, J.P. Lepeltier, and B. Marchal, in Ref. [13] pp. 294-301.
25. R.J. Elliott, in Ref. [11], pp. 142-155.
26. L.C. Evans, Trans. Amer. Math. Soc. (1983), No. 275, pp. 245-255.
27. W.H. Fleming, SIAM J. Control 9 (1971) pp. 473-517.
28. _____, Applied Math. Optimiz. 4 (1978) pp. 329-346.
29. _____, J. Optimiz. Theory Applic. (1983).
30. _____, SIAM J. Control Optimiz. 20 (1982) pp. 286-301.
31. _____, in Ref [12] pp. 131-141.

32. W.H. Fleming and S. Mitter, *Stochastics* 8 (1982) pp. 63-77.
33. W.H. Fleming and M. Nisio, submitted to *Nagoya Math. J.*
34. W.H. Fleming and E. Pardoux, *SIAM J. Control Optimiz.* 20 (1982) pp. 261-285.
35. B. Gaveau, *J. Funct. Anal.* 25 (1977) pp. 391-411.
36. F. Guerra and L.M. Morato, *Physical Review D* 15, Febr. 1983.
37. P.S. Hagan, R.E. Caflisch, and J.B. Keller, submitted to *Econometrica*.
38. O. Hijab, *IEEE Trans. Auto. Control*, Febr. 1983.
39. C.J. Holland, *Commun. Pure Appl. Math.* 30 (1977) pp. 755-765;
31 (1978) pp. 509-519.
40. I. Karatzas and V.E. Benes, *SIAM J. Control Optimiz.* 19 (1981) pp. 283-332.
41. P.R. Kumar and W. Lin, *IEEE Trans. Auto. Control* AC-27 (1982) pp. 765-774.
42. S.M. Lenhart and S.A. Belbas, *SIAM J. Appl. Math.* 1983.
43. J.P. Lepeltier and B. Marchal, *Stochastics* 2 (1979) pp. 243-286.
44. Y-C. Liao, Thesis, Brown Univ., 1982.
45. P-L. Lions, in Ref [12], pp. 199-215.
46. P-L. Lions and J-L. Menaldi, *SIAM J. Control Optimiz.* 20 (1982) pp. 58-81, 82-95.
47. D. Ludwig and C.L. Walters, *Ecological Modelling* 14 (1982), pp. 273-292.
48. P. Mandl, *Adv. Appl. Prob.* 6 (1974) pp. 40-60.
49. G. Mazziotto and J. Szpirglas, *Proc. 3rd IMA Conf. on Control Th.*, Academic Press 1981, pp. 567-586.
50. J-L. Menaldi, *SIAM J. Control Optimiz.* 18 (1980) pp. 722-739.
51. M. Nisio, in Ref [12] pp. 267-275.
52. H. Pragarauskas, in Ref [13] pp. 338-344.

53. R.W. Rishel, IEEE Trans. Auto. Control AC-22 (1977).
54. R.W. Rishel, IEEE Trans. Auto. Control AC-26 (1981) pp. 606-609.
55. M. Robin, Thèse d'état, Univ. de Paris IX, 1978.
56. M. Robin, SIAM J. Control Optimiz. 19 (1981) pp. 333-358.
57. D.D. Sheng, in Ref [12], pp. 353-375.
58. S.J. Sheu, Thesis, Brown Univ. 1982.
59. S. Stroock and S.R.S. Varadhan, Multidimensional diffusion processes, Springer Verlag, 1979.
60. N. Trudinger, Math. Research Report No.23-1982, Australian National University.
61. D. Vermes, in Ref [13] pp. 369-377.
62. A. Ju. Veretennikov, Math. USSR Sbornik 39 (1981) pp. 387-403.
63. B. Wittenmark and K.J. Åström, Proc. 21st IEEE Conf. on Decision and Control, Dec. 1982, pp. 33-37.
64. J. Zabczyk, in [13], pp. 144-154.